

AD-A253 258



2

A RAND NOTE

Thinking About Opponent Behavior in Crisis
and Conflict: A Generic Model for Analysis
and Group Discussion

Paul K. Davis, John Arquilla

DTIC
ELECTE
JUL 29 1992
S A

Document has been approved
for release and sale; its
distribution is unlimited.

RAND

92-20318



92 7 27 234

The research reported here was sponsored by the United States Air Force under Contract F49620-91-C-0003. Further information may be obtained from the Long Range Planning and Doctrine Division, Directorate of Plans, Hq USAF.

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of RAND do not necessarily reflect the opinions or policies of the sponsors of RAND research.

A RAND NOTE

N-3322-JS

Thinking About Opponent Behavior In Crisis and Conflict: A Generic Model for Analysis and Group Discussion

Paul K. Davis, John Arquilla

Prepared for the
Joint Staff

Accession For	
NTIS CRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DTIC QUALITY INSPECTED 3

RAND

Approved for public release; distribution unlimited

PREFACE

This Note, completed in May, 1991, was developed as part of a larger project on defense planning in the post-Cold War era, one task of which deals with understanding possible opponent behaviors in crisis and conflict. The project is supported by the Joint Staff and is being conducted in the International Security and Defense Strategy program of RAND's National Defense Research Institute (NDRI), a federally funded research and development center established by the Office of the Secretary of Defense and the Joint Staff. Comments are welcome. The authors' electronic addresses on Internet are Paul_Davis@rand.org and John_Arquilla@rand.org.

SUMMARY

This paper combines insights from strategic analysis, cognitive psychology, gaming, and artificial intelligence modeling to describe a theory and concrete methodology for thinking about the likely and possible reasoning of opponents before or during crisis and conflict. The methodology is intended for use in analysis and defense planning, especially planning for possible limited contingencies. We anticipate that the methodology can be employed in group discussions with policymakers and senior officers, not merely at the analyst-to-analyst level.

The basic framework is general (i.e., we do not build separate Russian, Arab, Latin, and Oriental models). Many cultural factors can be brought to bear in specific applications, but we do not believe, for example, that it is useful to assume that all or most Arab leaders think alike.

A fundamental tenet of the approach is that at least two semiformal models of the opponent should be developed and carried along through analysis and decisionmaking. Many opponent models are imaginable, but in any given instance, many of them will be obviously unsuitable and two or at most three will suffice to improve greatly the quality of discussion by forcing participants to confront the possibility of very different opponent mindsets. The multiple-model approach is related to but goes beyond the familiar but often ineffective method of creating a devil's advocate. It would also be implemented not as a last-minute bow to openmindedness, with the deck stacked against the devil's advocate, but as a structural change in analytic and group-discussion procedure.

Psychologically, forcing people to confront multiple models mitigates tendencies toward overconfidence in prediction, failure to hedge, and failure actively to find ways to determine and affect the opponent's reasoning before it is too late. It also legitimizes alternative views. There is a strong body of empirical evidence to support the importance of this approach, much of it associated with effects of *framing, anchoring, attributional inference, and groupthink*. To build opponent models, regional specialists should be teamed with strategists and model-oriented analysts; they should then be required to develop the alternative models and proscribed from slipping toward a best-estimate consensus.

The opponent models at issue here are descriptive rather than prescriptive. They exhibit what we call *limited rationality*, which is quite different from the idealized economic rationality or rational-analytic style of thinking emphasized in standard decision analysis and the theory of games. A leader with limited rationality will make decisions that bear a

reasonable relationship to his objectives and values, but they may be flawed by misperceptions, miscalculations, cognitive biases, and the tendency to either accept more or fewer risks than complete "rationality" would call for—and, in the risk-accepting case, to focus on maximizing the likelihood of complete success rather than maximizing the expected value of some utility function. Our models can reflect much, and perhaps most, of what often is passed off as irrational behavior on the part of aggressors. They also can reproduce a wide range of psychological phenomena that are seldom described analytically.

A key assumption in our approach, based on behavioral science's *prospect theory*, is that possible opponents are likely to become increasingly and "unreasonably" risk accepting as they become emotionally more dissatisfied with their current situation and trends. By contrast, opponents who are more comfortable with the status quo are likely to be more risk averse than rational-analytic theories might predict. Other factors affecting risk taking in our model are the degree to which the decisionmaker perceives himself to be in *control of the initiative* and the historically common analytic bias of *underestimating nonimmediate power*—e.g., underestimating the ability and willingness of the United States to deploy forces from afar and underestimating the long-term impact of those forces.

Our approach frames decision issues in a natural way. Instead of formulating utility functions and calculating expected utility, we describe decisionmaking as a process of balancing different considerations *qualitatively*—considerations that include, for each option assessed, best-estimate (likely), best-case, and worst-case outcomes. Thus, in our model's imagery, the decisionmaker might conclude his deliberations by saying, for example:

We agree that the current situation is intolerable and getting no better. Second, we see that prospects are not great—we could fail and chances of success are only marginal. And, yes, the risks are high. However, there is also the temporary potential for complete success—one that would bring honor to our cause and to all of us here. There have been other options discussed that would probably result in some improvements. None, however, would change the situation fundamentally, and an opportunity would have been lost. My decision, therefore, is to proceed. Military operations will begin at dawn.

This imaginary decisionmaker appears to exhibit limited rationality, even though (if one reads between the lines) he may be leading his people into disaster. Note also that he reasons in terms of qualitative concepts such as prospects, risks, and opportunities, somehow folding them together to reach a judgment. In doing so, he is not conducting a mathematical analysis. Instead, he is using heuristic reasoning, as all of us do because of the problems of *bounded rationality*. He is also applying thresholds implicitly. And he is making his judgments and comparisons within an implicit mental *frame* and relative to certain

anchor points. Our models make the assumptions more explicit. And, as do people, our models depend on hierarchies of variables. Thus, if an assessment of risks enters into a top-level decision, then the logic for the assessment of risks constitutes a second, lower-level decision dependent on additional variables. The second-level model logic might be equivalent to saying "Well, since the economic and military risks are low, and the political risks are high only in the short term, we can conclude that risks are low overall."

To represent this type of reasoning for purposes of analysis and group discussion, we use diagrammatic techniques and decision tables that have proven effective in RAND efforts over the last several years. Simplified versions lend themselves well to small brainstorming sessions at a blackboard or to group discussions with large-screen interactive graphics.

Even with the best modeling methods and regional specialists, any decision modeling is likely to be highly uncertain. However, the process of modeling the opponent and understanding the two or several ways he *may* be viewing the situation can nonetheless be very useful in developing strategy. Our objective in the endeavor is less to predict confidently his thinking and behavior—from an aloof position of disinterest—than to understand how the opponent *may* be reasoning, how we might affect his reasoning, and how we should prepare generally—hedging our bets and acting proactively as appropriate. Our approach based on multiple opponent models appears to have considerable potential for that purpose, so long as the multiple models are maintained and taken seriously as a basic requirement of the approach. We note in this connection that the mindsets of real people are not only uncertain, but also subject to discontinuous change when they suddenly "see things differently." Thus, even the concept of "the correct" model is flawed.

In a companion paper (Davis and Arquilla, 1991) we discuss our experimental use of the approach during the crisis that preceded the U.S. offensive operations against Iraq. In ongoing work we are drawing on other historical crises to evaluate the approach. The government should consider experimental gaming using the approach to think about potential future crises. Although we are optimistic about the approach's potential, there is much to be learned and added.

ACKNOWLEDGMENTS

We appreciate the many helpful suggestions of numerous RAND colleagues, especially Robert MacCoun, who reviewed the draft. We also appreciate discussions with Alexander George of Stanford University and the U.S. Institute for Peace, Lee Ross of Stanford University, Bruce Bueno De Mesquita of the Hoover Institution, and Daniel Kahneman of the University of California at Berkeley. They, of course, bear no responsibility for our arguments here.

CONTENTS

PREFACE	iii
SUMMARY	v
ACKNOWLEDGMENTS	ix
FIGURES AND TABLES	xiii
Section	
I. INTRODUCTION	1
Objectives	1
Motivation for Study	1
II. GENERIC FEATURES OF THE OPPONENT MODELING APPROACH	6
Prior Work	6
Broad Precepts	6
Influence Diagrams	16
Decision Tables	17
III. APPLYING THE METHODOLOGY	21
Background Staff Work	21
Group Discussion	25
Analysis	25
Amending One's Own Strategy	26
Appendix	
A. CONDITIONAL RATIONALITY	29
B. COMPARISON WITH STANDARD DECISION ANALYSIS	31
BIBLIOGRAPHY	35

FIGURES

2.1	An Idealized Process Model of Decisionmaking at a Snapshot in Time	11
2.2	Risk Taking as a Function of Perceived Situation	12
2.3	A Hierarchy of Variables	15
2.4	One Possible Hierarchical Depiction of Saddam's Value Structure in 1990	16
2.5	An Illustrative Influence Diagram	17
3.1	An Exercise in Humility: Past Shocks	22
B.1	Special Case in Which Outcomes and Probabilities Map Easily into Constructs of Current Model	32

TABLES

1.1	Special Problems in Limited Contingencies	2
2.1	An Illustrative Judgment Table	19
2.2	Efficient Representation of Same Judgment Table	20
2.3	An Options-Comparison Table	20
B.1	A Possible Risk Taker's Assessment of Options	33

I. INTRODUCTION

OBJECTIVES

This Note is part of a study of peacetime contingency planning and crisis decisionmaking, the objectives of which are (a) to develop a conceptual approach for thinking about the likely and possible reasoning of opponents before or during crisis and conflict, (b) to evaluate the approach against historical cases (including the recent crisis in Iraq), and (c) to refine techniques for applying the approach in both analysis and group discussions. Here we are concerned primarily with (a) and (c). In a companion paper we apply the approach to the decisionmaking of Saddam Hussein prior to the beginning of U.S. offensive operations in January, 1991 (Davis and Arquilla, 1991).

MOTIVATION FOR STUDY

New Focus for Defense Planning

As the Cold War came to an end in 1989 and early 1990, it became clear that defense planning should increasingly focus on deterring or coping with a wide variety of possible contingencies worldwide, rather than continuing to dwell on all-out global war with the Soviet Union. The contingencies in question might vary from political crises, with some movement of military forces for signalling, to sizable wars.

Typically, the best way to deal with a contingency is to avoid it—i.e., to head off a crisis before it occurs. The next best way is to defuse the crisis promptly. A much less desirable way is to engage in a brief, limited, but successful shooting war; even less desirable is a longer and larger war, even if successful. The worst is to have a long, substantial, and unsuccessful war such as that in Vietnam. In some cases, of course, there are objectives beyond avoiding or bringing an early end to violence; war may even be preferable to the alternatives.¹ By and large, however, we prefer that deterrence work.

To face up to this rank ordering, defense planning must include explicit planning not only for warfighting, but also for general deterrence in peacetime, specialized deterrence actions at the outset of crisis, and limited military actions to achieve successful crisis outcomes quickly. Such limited actions, however, must be taken with due regard for the possibility that deterrence may fail and substantial combat occur.

¹Anticipating some later discussion, note that the emphasis here on avoiding or defusing crisis is characteristic of a nation that largely enjoys the status quo. We should not expect other actors in the international scene, particularly the disenfranchised, to have similar values.

The need to reflect all these factors in defense planning has been recognized and emphasized since late 1989 by the Joint Staff, and the crisis and war in Iraq has strongly confirmed it.²

Challenges in Thinking About Limited Contingencies

It is one thing to posit paying more attention to limited contingencies and changing the way one plans; it is another to do it. There are many challenges associated with limited contingencies that do not arise in considering all-out global war with another superpower. Table 1.1 indicates some of them.

Table 1.1
Special Problems in Limited Contingencies

-
1. *Knowledge.* The sides may not understand each other's interests, objectives, strategies, or capabilities.
 2. *Instruments of Persuasion.* The various instruments of action may be maldeployed, undefined, poorly prepared, or ill coordinated—especially across government agencies and across governments that need to act in concert.
 3. *Communications.* The sides may have physical difficulty communicating, or communicating *effectively*, because of poor personal contacts and cognitive disconnects that cause signals or even intendedly straightforward messages to be misperceived. Tacit bargaining may be difficult.
 4. *Decisionmaking.* U.S. decisionmaking may suffer by starting late—i.e., until the crisis “breaks,” the trouble spot may not even have been high in the consciousness of key officials.

The “rationality” of the opponent’s decisionmaking may suffer from lack of staff work and free debate with senior advisors, especially in totalitarian regimes having a single dominant figure.

All participants’ decisions may be influenced by personal, political, and organizational factors.
 5. *Command and Control.* The opponent’s command and control system may be technically and organizationally weak. The opponent’s forces may be undisciplined, untrained, and unclear about how to behave, making military “accidents” likely.
 6. *Values.* Ultimately, the sides may *not* share common values. In particular, the opponent may not share the goal of avoiding crisis or having the crisis defused early and without violence—especially if that means a return to the status quo ante.

Further, the sides’ or leaders’ value systems may be very different regarding what might constitute a “fair” or “honorable” outcome.
-

²For related discussion, see Davis (1989a) and Winnefeld and Shlapak (1990).

To better appreciate these points, consider them in terms of the confrontation between superpowers during the Cold War. The United States and Soviet Union have had extensive knowledge about each other, including military capabilities, for decades. The instruments of deterrence and warfighting have been in place or quickly available in crisis (especially the nuclear deterrent forces). Communications have not been a limiting factor in decades, and both sides have been sensitive to its importance. Both countries have been acutely concerned about possible crises, and their leaders have regularly turned attention early to any superpower crisis that might be developing. Both nations' governments are large, mature organizations with a tradition of distributed responsibilities and group discussion (when considered in comparison with, say, the governments of Qaddafi, Saddam Hussein, or Kim Il Sung), which tends to dampen some of the more serious forms of nonrational reasoning, although introducing other problems.³ Both countries' leaders have—to greater or lesser degree (least under Stalin)—recognized that the threat of general nuclear war requires a form of thinking that transcends concerns about personal prestige, marginal gains, and reaching decisions internally that represent compromises across organizations. Both countries have had firm control over their military forces, with sophisticated worldwide communications. And, finally, both nations' political leaders have shared values that meant a nuclear crisis would not be viewed as a zero-sum game. In particular, both sides' political leaders have believed that a superpower war might very well become nuclear, that nuclear war would be extremely difficult to control, that general nuclear war would be an unmitigated catastrophe for all mankind, and that great pains should therefore be taken to avoid not only nuclear crises but superpower military crises more generally.⁴

Other Challenges for Contingency Planning

Planning for limited contingencies also raises some serious tensions. In some cases, there will be conflicts between the actions one would take if maximizing the likelihood of

³This distinction between Soviet and third-world-dictatorship command-control systems has been articulated by colleague Paul Bracken. Staff processes, including group discussion, often increase the options considered and improve the quality of information available to decisionmakers. On the other hand, there is strong empirical evidence of various *groupthink* problems (Janis, 1982) such as *polarisation*, in which group discussion tends to reinforce unduly the initial preference or judgment of the group. (See reviews by Myers and Lamm, 1976, and Isenberg, 1986; see also Whyte, 1989, which reinterprets some of the evidence on groupthink in terms of framing problems.)

⁴During the Cold War, of course, there was a great deal of controversy about the degree to which the sides shared these views. In the late 1970s and 1980s, many Americans viewed Soviet military capabilities and military doctrine with alarm, particularly because there seemed to be an attitude that nuclear wars could be fought and won. Ironically, Soviets had similar views about American attitudes in the late 1950s and early 1980s. See Davis and Stan (1984) for a mid-1980s discussion that points out many differences of mindset while concluding that the sides' leaders agreed firmly on the basics mentioned in the text.

defusing a crisis early and the actions one would take if maximizing the likelihood of successful warfighting. A small "tripwire" force, if deployed early, might deter. If it did not, however, that force might be lost in the initial hours of combat. Delaying deployment until large forces could be inserted might reduce the likelihood of such losses, but might sacrifice the opportunity to defuse the crisis without war.

Yet another challenge for the new era's defense planning is that it is essential, not merely virtuous, to develop strategies and exercises combining political, military, and economic instruments. A corollary is that military planning should recognize the near certainty of numerous constraints and of possible requests for actions that might seem inappropriate from a warfighter's perspective.⁵

Finally, contingency planning seems to require worrying about the reasoning and behavior of possible opponents during crisis and conflict. It is this challenge that is of most concern in this paper.

The Need for Opponent Modeling

Clearly, if one is to have strategies for peacetime deterrence and for enhancing deterrence in crisis, one must have some sense of how specific opponents would react to different measures.⁶ We need a theory of deterrence and a larger theory of strategy for limited contingencies. Further, we need a methodology for *applying* these theories to specific opponents in specific contexts. This will include building appropriate models.

There already exists an extensive literature on deterrence and the strategy of conflict, much of which is as sound now as when it was first developed, but the vast majority of that literature is limited by one or more of the following restrictions of scope: (a) superpower nuclear crises; (b) rational-actor reasoning; (c) emphasis on purely technical and quantitative considerations (e.g., the counterforce-exchange problem); (d) *descriptive* accounts of why deterrence has not succeeded historically, because of various cognitive problems such as misperception and flawed mental processing of information; and (e) general prescriptive advice for improving U.S. crisis decisionmaking *processes* (e.g., admonitions to actively seek

⁵Yet another factor in the new defense planning is the need to give priority to maintaining certain specialized forms of support forces in active status, even at the expense of some active-component combat forces. Such support forces can be critical in the early stages of crisis action (see, e.g., Winnefeld and Shlapak, 1990).

⁶Actually, one also needs to understand the reasoning and behavior of other relevant nations, including potential allies and the opponent's potential allies. We shall not discuss these matters in this paper, however.

out alternative strategies). There has been remarkably little done on opponent modeling and methodology for adapting strategy to particular opponents in crisis and conflict.⁷

All of this, then, was the background motivating the current study. We were more optimistic than conventional wisdom dictates about being able to make progress, primarily because of the senior author's earlier experience modeling U.S. and Soviet decisionmaking in possible nuclear crises and the consequences of that work for thinking about issues such as flexible-response doctrine and improving first-strike stability. Nonetheless, we began the study with considerable trepidation.

⁷Even in the heavily studied domain of nuclear strategy there has been little formal opponent modeling, except through game theory focusing primarily on quantitative measures of military destructive capability. For exceptions, see Davis, Bankes, and Kahan (1986) and Davis (1989b), which model U.S. and Soviet decisionmaking in crises at or near the nuclear threshold. These emphasize behavioral considerations and human values rather than the standard calculations of strategic nuclear analysis.

II. GENERIC FEATURES OF THE OPPONENT MODELING APPROACH

PRIOR WORK

The approach described in what follows draws on prior work with artificial intelligence concepts (e.g., Davis, Bankes, and Kahan, 1986, and Davis, 1989c), which resulted in large and complex knowledge-based computer models of the United States and Soviet Union in large-scale crises. These models have been used in a man-machine system for analytic war gaming (the RAND Strategy Assessment System, or RSAS), as well as in discrete studies. By contrast, the approach taken here is simpler and more suitable for use in a variety of contingency applications. Further, the approach taken here is strongly influenced by the need for interactive group discussion in both defense planning and periods of crisis.

BROAD PRECEPTS

Multiple Models and Framing Effects

A basic concept derived from the earlier work is the importance of developing and carrying along at least two alternative models of the opponent. This concept is in direct contrast with the more standard approach of seeking consensus on a best-estimate image, which we believe is manifestly wrong headed in view of the uncertainties and complexities that are usually present, as well as the fact that real people (and groups) appear to shift back and forth among alternative mindsets.⁸ The importance of the multiple-model approach is discussed below and illustrated in our companion paper (Davis and Arquilla, 1991), which deals with alternative images of Saddam Hussein.

The need for alternative images is closely related to a central concept in cognitive psychology, notably the importance of *framing* and the related issue of *anchor points*. When people evaluate options (and decide whether to take or avoid risks), their conclusions are highly dependent on how the problem is presented (i.e., framed). Further, they tend to compare options relative to some standard (anchor point) such as *status quo* on the one hand or a *goal state* on the other. Depending on the framing and anchor point, they may even reverse their *relative* preference for options, even though doing so violates a fundamental transitivity axiom of rational thinking (see particularly Tversky and Kahneman, 1981; see also Karni, 1990, and Lichtenstein and Slovic, 1971). This phenomenon is familiar to all of us in our everyday lives. Strategy should seek to affect the opponent's framing and

⁸The flip-flopping of mental frames is discussed in Davis (1989b), Kull (1988), and Kanwisher (1989).

anchoring, and should seek also to assure that our own framing and anchoring are sound. Using multiple models and alternative representations of the world can increase the open-mindedness that is necessary in doing both.⁹

An issue here is how many models are necessary. Could there not be need for dozens of alternatives? We believe not. We believe that a half dozen generic models will probably turn out to have broad applicability, and that in a particular instance it will be evident that only two or three of them could possibly apply. For example, to understand Japanese decisionmaking before World War II or Argentine decisionmaking before the Falklands War, one should consider models allowing for strong organizational factors. However, such models would seem to be clearly inappropriate in thinking about Saddam Hussein. There is much research yet to be done on these matters, but part of the approach we recommend is to winnow down the possible models to a workable level (two or at most three). Importantly here, we do *not* believe that it is necessary to fine-tune the models used or to choose the very best of some comparable models. In our experience with human war gaming and strategic planning exercises, as well as in our more limited experience with opponent modeling, there is a great benefit to opening minds to the possibility of even a second way of looking at a problem. Once that occurs, discussion and thinking shift toward facing up to uncertainty and developing ways to hedge and adapt as a function of circumstances. Even if the models used prove to be not quite correct, there is in practice a good chance that relevant strategies will have been surfaced and considered, and that "the gameboard" will be understood.

Culture-Free Structure

Our framework is intended to be general enough to be useful for thinking about highly diverse opponents. This approach is in contrast to one that would involve different models for "the Arab mind," "the Latin mind," and so on. Differences in culture and the like definitely enter into our approach (we are by no means mirror imaging!), but they enter via the specifics of an application, not at the level of structure or fundamental concepts. The reason here is straightforward: we believe that one can find in all cultures a wide range of reasoning styles, and it would be a first-order mistake to assume otherwise. Just mentioning the names of Saddam Hussein, Anwar Sadat, and King Fahd of Saudi Arabia should

⁹There is a substantial literature on all of these matters. T. C. Chamberlain argued the importance of multiple hypotheses some decades ago (see discussion in the classic article by Platt [1964]). In the behavioral literature, there are numerous studies demonstrating serious cognitive biases due to the failure to consider alternative hypotheses. See, e.g., Fischhoff (1982), Hawkins and Hastie (1990), Ross and Anderson (1982), and Tversky and Kahneman (1982). See also the highly readable review monograph by Ross and Nisbett (1990) and recent papers about the origins of overconfidence (Vallone, Griffin, Lin, and Ross, 1990, and Dunning, Griffin, Milojkovic, and Ross, 1990).

demonstrate the folly of thinking in terms of *the Arab mind* except for rhetorical purposes. Similarly, it would be foolish to believe that only the Japanese, with their particular culture, would allow fatalism about the inevitability of war to drive them into a surprise attack such as that at Pearl Harbor. And, to give yet another example, does anyone really believe that the long view of Ho Chi Minh was unique to Oriental culture? Whoever believes it was might wish to review the strategy of Peru's *Shining Path* movement. In summary, we certainly believe that country-specific and leader-specific considerations of culture and context should be reflected, but we do not believe in framing opponent models in terms such as "the Arab mind."

Limited Rationality

Perhaps the most important decision about a modeling approach involves the issue of "rationality." There is a view that many historical protagonists have been so irrational as to preclude attempts to understand them. Hitler, Stalin, Qaddafi, Khomeni, and Saddam Hussein have all been characterized by many as irrational. Our approach, however, is based on assuming what we call *limited rationality*, by which we mean that the actual decisions and behavior of leaders in crisis bear a "reasonable" relationship to objectives and values.¹⁰ Decisions and behavior are not seen as random, as they might be if the decisionmakers suffered from the effects of hallucinogenic drugs or organic mental illnesses. Instead, decisions and behavior follow a relatively straightforward logic from assumptions, perceptions, objectives, and values (all of which may be much less straightforward). Limited rationality does not apply to all relevant leaders, but we believe it applies well enough to be quite useful for the vast majority.¹¹ Even most of the decisions of Hitler, who *did* suffer from physical and drug-induced problems, as well as from psychological problems, seem to have been more regular and understandable (however despicable) than is sometimes assumed. To put it differently, Hitler's "style" was understandable.

Limited Rationality vs Economic Rationality and Rational Actors

We distinguish limited rationality from what is often associated with rational-actor models. In particular, we do not assume that the decisionmaker acts to maximize an explicit or implicit utility function (economic rationality or rational-analytic reasoning). To the contrary, decisionmakers are subject to the nonrational aspects of human decisionmaking

¹⁰See Appendix A for a more extensive discussion of different concepts of rationality.

¹¹As Thomas Schelling noted 30 years ago, decisionmakers are not characterized by location on a single line extending from rational to irrational. Instead, "rationality is a collection of attributes, and departures from complete rationality may be in many different directions" (Schelling, 1980:16).

that we all share biologically, which are more dominant in some kinds of decisionmaking than in others.¹² These include miscalculating the balance of positives and negatives in considering options, misestimating probabilities, ignoring low-probability outcomes, preferentially focusing on good news (or bad news), and so on. Further, the decisionmakers we have in mind are all subject to *bounded rationality*, as described in Simon (1982) and his much earlier works. Even with a staff of Ph.D. decision theorists to help them, their decisions would be strongly affected by uncertainties about the facts of the situation and the likelihoods of various events. They are also subject to certain analytic and behavioral biases and may well use what are sometimes called *cognitive heuristics*, which often turn out to be counterproductive (Kahneman, Slovic, and Tversky, 1982, and Kanwisher, 1989).¹³

Our approach of limited rationality also recognizes that *we cannot expect to predict decisions accurately and consistently, no matter how much effort we put into understanding the individuals involved and the information they are likely to have before them*. There are random factors at work—for example, who gives the last advice and which ideas are on the table when decisions must suddenly be made quickly, factors described in work by James March and others (e.g., March and Weissinger-Baylon, 1986). Furthermore, it is not possible in principle to understand everything about the opponent leaders, their staffs, and their information. Thus, our approach recognizes uncertainty (through the use of multiple models and other mechanisms) and makes no attempt at fine-tuning. Consistent with this, contingency planning or crisis planning should not place too much of a burden on subtle signals or correctly guessing which opponent model is “correct.” Instead, it should seek to shape circumstances so as to maximize the likelihood of opponent decisions consistent with our own objectives.

Unilateral Actors vs Organizational Actors

Our approach appears at first glance to assume a “unilateral actor,” and to ignore the many interesting and crucial roles that organizational factors may play in a government’s decision, as discussed in various classic books (e.g., Allison, 1971, Steinbruner, 1974, and Arrow, 1963). The reality of our approach is more complicated. First, however, let us

¹²There is a rich and growing literature on the cognitive limitations of human decisionmakers. Some classic references include Jervis (1976), Axelrod (1976), March and Weissinger-Baylon (1986), Janis and Mann (1977), George and Smoke (1974), and Kahneman, Slovic, and Tversky (1982). Davis (1989b) describes how such factors might play in a nuclear crisis. Holsti (1989) provides a good review and bibliography on crisis decisionmaking. For discussion of group behavior, see Janis (1982), Myers and Lamm (1976), and Whyte (1989). Ross and Nisbett (1990) is an excellent and readable overview of the field, with an extensive bibliography.

¹³We concern ourselves with *cognitive biases* rather than what the literature refers to as *motivated biases*.

observe that many of the opponents of most interest in defense planning have to a substantial extent behaved as though they were unilateral actors, at least on the strategic level. Hitler, Mussolini, Saddam Hussein, and Stalin come to mind here. It is always a matter of degree, of course, since even strong leaders are influenced and constrained by organizational and political realities (e.g., Franklin Roosevelt's policies during the 1930s were strongly constrained). Nonetheless, there is reason to believe that a unilateral-actor model will often be appropriate in U.S. contingency planning.¹⁴

This said, we do not wish our approach to be constrained by unilateral-actor assumptions. Thus, in applying the approach, we should be altogether willing to bring organizational and political factors to bear as seems appropriate. For example, in considering how a particular opponent might assess the consequences of one of our actions, we might ascribe to that opponent the views of his Navy if we believed the Navy would be organizationally dominant with respect to his decisions on the matter at hand. Also, we may develop alternative opponent models to reflect different views about which influences are dominant.

A Virtual Process Model of Decisionmaking

A second basic feature of our approach is the assumption that decisions are made *as though* they were the result of a process that includes situation assessment, adjustment of assumptions, framing and assessment of options, and follow-up analysis to see if the favored option can be implemented (Davis, Bankes, and Kahan, 1986). Figure 2.1 shows an idealized decision process at one point in time. Actual decisionmaking is often distinctly nonlinear and complex, but in many cases the results are the same *as though* the process had been linear. Further, one of the principal functions of staff work and advisors is to improve the degree to which decisionmaking is consistent with the idealized model of decisionmaking. In this paper we use only some of the structure suggested in Fig. 2.1, but the image of this structure underlies much of what we do.¹⁵

¹⁴See also the related discussion in Bueno De Mesquita (1981).

¹⁵In prior work (e.g., Davis, Bankes, and Kahan, 1986), a computerized artificial intelligence model using the idealized process of Fig. 2.1 was actually given greater sophistication by allowing some feedback (iteration) within a decision process. In addition, there was iteration as a result of decisions being revisited, with some of the earlier judgments being used in the revisiting. That is, the computer models needed to have some memory of past decisions to be realistic. The absolute minimum memory was that of the strategy previously decided upon.

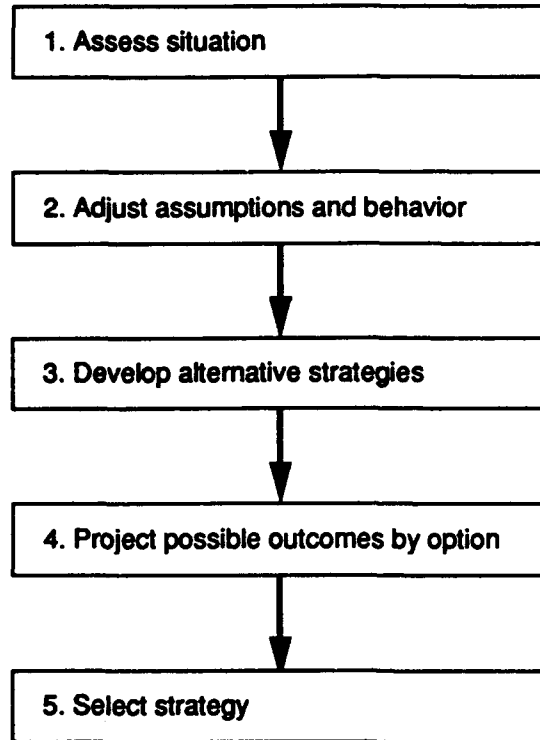


Fig. 2.1—An Idealized Process Model of Decisionmaking at a Snapshot in Time

Considering Upside and Downside Potential

Another key element of our approach is to assume that decisions are *strongly* affected not only by a “best estimate” of outcome for each option considered, but also by an assessment of upside potential, downside risks, or both. This assumption may seem obvious, but it is common in analytic work to focus on maximizing expected outcome. Even sophisticated game theory methods are almost always applied with assumptions such as that both protagonists are trying to maximize the expected value of their utility function (see, e.g., Schelling, 1980, Brams, 1985, and references therein). By contrast, real people typically are more sensitive to either risks or upside potential than such theories predict. Figure 2.2 shows this type of effect schematically and indicates the empirical result (called *prospect theory*) that people’s behavior shifts from risk avoidance to risk acceptance as they become more and more dissatisfied with their current lot—either as a result of changes in the situation or as a result of changes in the way they *frame* the problem (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981). Game theory methods usually assume people make decisions that are “economically rational” (the dashed horizontal line).

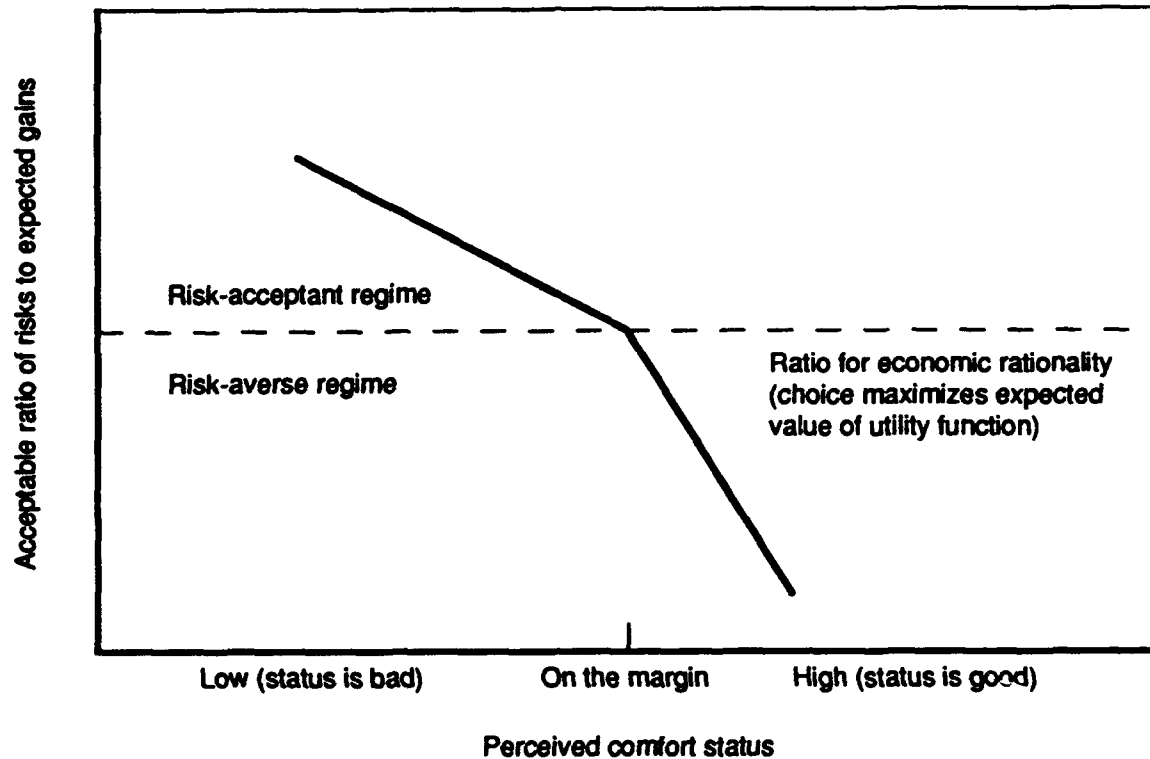


Fig. 2.2—Risk Taking as a Function of Perceived Situation

Although the decision-theory literature depends heavily on empirical research with common people or students, "nonrational" treatment of risks and opportunities is also visible historically at the level of national leaders and generals. We are accustomed to talking about Adolph Hitler and Ariel Sharon as being risk taking and about Neville Chamberlain and George McClellan as being very risk averse. As will become clearer later, our approach seeks to highlight the factors allowing us to make these distinctions, rather than burying them in mathematics.

Situational and Behavioral Influences on Framing and Judgment

Although the framework of our decisionmaking model may appear highly rational, structured, and universal, details matter. In particular, every step of the decisionmaking process is strongly affected by the decisionmaker's view of the world, perception of facts, and so on. Further, he may do a poor job in folding together information that may be available to him—because of biases, mindsets, and a variety of other factors (e.g., see Fig. 1.1 of Holsti, 1989, or somewhat similar arguments in Davis and Wolf, 1991).

At the behavioral level it is particularly important to understand, for the particular individual whose decisionmaking is to be modeled, how he would see:

- The relative value of different outcomes (taking into account his concepts of fairness as well as his personal ambitions and fears)
- The likely behavior of other national actors (including not only the United States, but also his own regional neighbors)

Similarly, it is important to consider whether the individual seems likely to be affected by psychological factors such as:

- Paranoia
- Megalomania
- Desperation
- A sense of relative deprivation
- Specific fears, which might or might not be paranoid

These factors, which would be addressed in any psychological profile drawn up by regional experts or intelligence-community psychological experts, can strongly affect reasoning in the model (Davis and Arquilla, 1991).

Another important factor is the degree to which the opponent, especially a would-be aggressor, believes he possesses the initiative. Generally speaking, the ability to select the timing, scope, and intensity of an aggressive act gives the initiator a substantial latitude among options. At the same time, the task of any potential defender is exceedingly complicated by not having the initiative, so much so that deterrent postures may not be formulated or implemented against all of the aggressor's viable options. Further, the aggressor may, for complex political and social reasons, enjoy escalation control—being able to increase or decrease the level of conflict as appropriate. This is a classic difficulty in dealing with guerrilla organizations. One could also see such factors at work in Sadat's behavior before the Yom Kippur War and perhaps in Saddam Hussein's behavior relative to Iran over a period of many years.¹⁶

Yet another behavioral factor worth mentioning here is that many of the historical initiators of war have clearly shared an "analytic bias" in which they had difficulty comprehending adequately and emotionally certain forms of power that were not immediate

¹⁶See George and Smoke (1974:520-532) for related discussion of initiation theory in which the authors mention the multiplicity of options available to the aggressor as an important factor.

and familiar. This has long been a problem with respect to continental powers underappreciating the capability of maritime powers to thwart their ambitions. It may also have been a factor in Saddam Hussein's underappreciating both U.S. deployment capability and the potential of U.S. airpower, as evidenced by his public message to President Mubarek on January 2, 1991, that "we are not intimidated by navies."¹⁷

Risk Acceptance and "Going for the Gold Ring"

A particularly important aspect of decisionmaking "style" involves degree of risk acceptance and the degree to which the decisionmaker is fixated on achieving "full" success rather than incremental success. Whereas ordinary decision analysis assumes decisionmakers seek to maximize expected utility as described by a "reasonable" utility function—assumptions that favor worrying about risks and accepting partial successes—history is replete with figures who simply do not fit that mold. Indeed, we would argue that many conquerors and great generals have not done so. In terms of game theory, we would argue that such figures seek, within limits, to maximize the likelihood of glorious success.

Individual leaders are not, however, equally risk acceptant in all circumstances. To the contrary, we believe that a behavior of risk acceptance and "going for the gold ring" is most likely when (a) the current situation and trends are regarded *emotionally* as "intolerable," (b) the leader believes he controls the initiative, and (c) there are exacerbating psychological factors such as megalomania, paranoia, and a sense of relative deprivation. Our assumptions here correspond on the one hand to everyday experience with children, spouses, and schoolyard bullies. They also are consistent with our reading of the empirical behavioral literature on individual choice. Scaling this experience up to the behavior of national leaders in international crises is a leap of faith (criticized, for example, by Kull [1988]), but one we believe is consistent with experience. Certainly, if we are wrong, so also are historians, since historians routinely describe decisions with prefacing phrases such as "flushed by success and confident of his ability to mold events."

Hierarchical Structuring of Key Variables

Any serious model of decisionmaking must eventually consider a vast number of variables. However, for both analytic convenience and because it corresponds to how higher-level people actually make decisions, we organize our model hierarchically (Davis, Bankes,

¹⁷See Chapter 3 of Arquilla (1991) for a more extended discussion with historical evidence.

and Kahan, 1986). Thus, the decisionmaker in our model tries to reason at the level of "Given that the situation now is bleak and that by pursuing the option in question we would have good prospects for an acceptable outcome, some opportunity for a terrific outcome, and only moderate risks, we will go ahead." In this depiction, the final decision is based on looking at the variables situation, prospects, opportunity, and risks, each of which has some qualitative values such as good, moderate, or bad (or, e.g., bleak). However, if one asks why the situation is bleak, the answer must involve a whole series of additional variables. For example: "Given that even though the current economic picture is OK, the trends are bad, and that the political situation is poor, we have to consider the overall situation to be bad."

Figure 2.3 shows another example of hierarchical variables, this one suggesting that in evaluating an option, the leader in question might look separately at economic, political, and personal prospects, and within each of those at both short-term and long-term prospects. He would then fold these considerations together in reaching an overall judgment. The variables and their relative weights are specific to decisionmakers and contexts. Figure 2.4 suggests, using thickness of arrows, how Saddam Hussein might have weighed different considerations in contemplating different possible outcomes of the crisis as of August, 1990. Note the prominence given to Saddam's personal status.

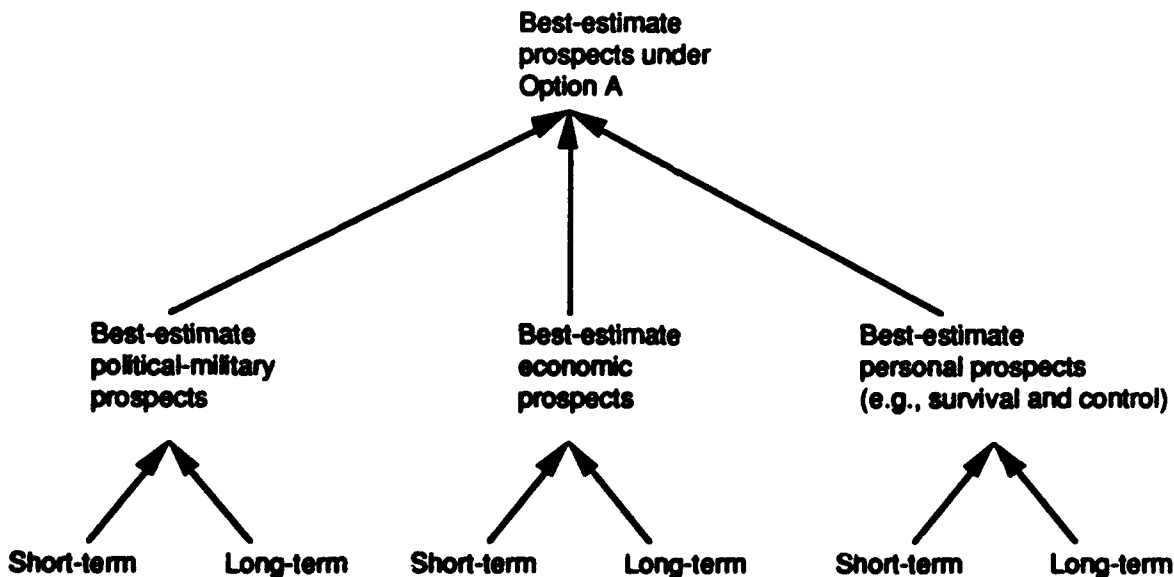


Fig. 2.3—A Hierarchy of Variables

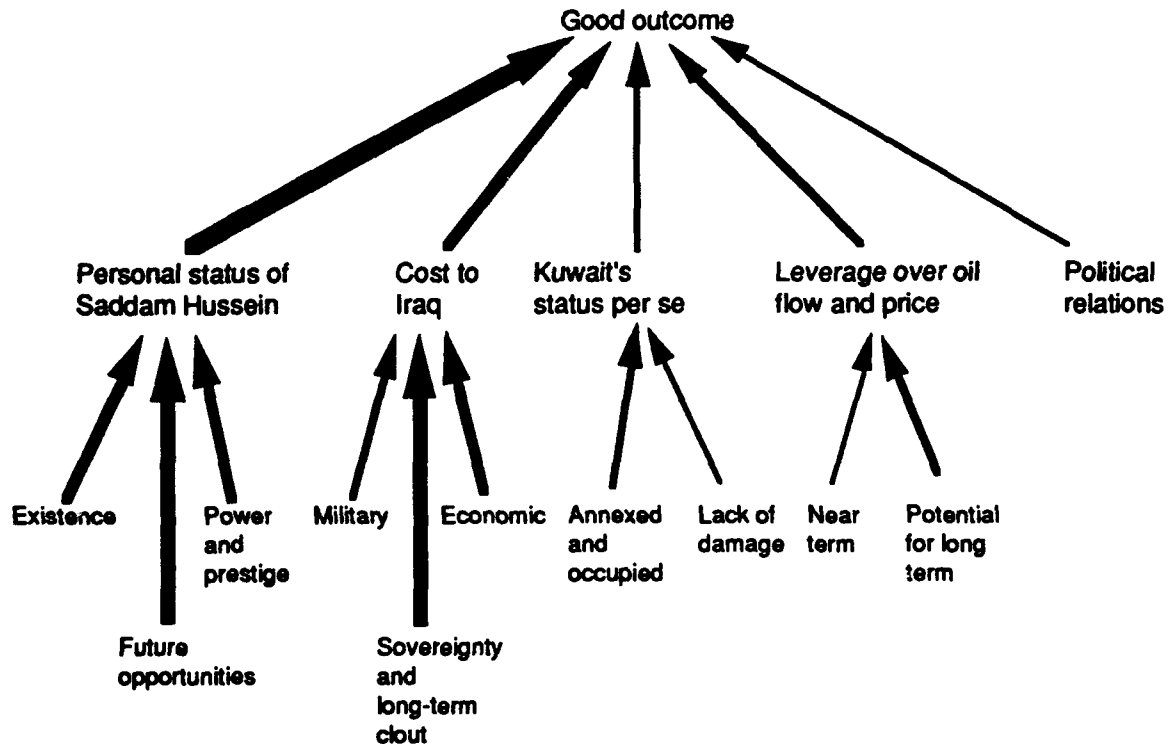


Fig. 2.4—One Possible Hierarchical Depiction of Saddam's Value Structure in 1990

INFLUENCE DIAGRAMS

Influence diagrams and the related *cognitive maps* are important tools in developing a sense for the key variables and for communicating images of real or perceived cause-effect relationships. Influence diagrams come in many different forms and probably have many inventors, but are usually associated with MIT's Jay Forrester and the Systems Dynamics method; cognitive maps were used by Robert Axelrod in his book *The Structure of Decision* (Axelrod, 1976). We have used our own versions of both for some years and have found them quite valuable. Figure 2.5 provides one example. The convention here is that if an arrow connects two items, the first item (or an increase in it) tends to cause the second (or an increase in it), unless there is a negative sign affixed to the arrow, in which case the first item tends to work against or decrease the second. In the example, we see the tension mentioned earlier: deploying a tripwire force early in crisis may upset the opponent's calculations, regain the initiative, and generally increase chances for crisis resolution. However, it may also cause casualties if war occurs, and preclude the option of staying out of the conflict.

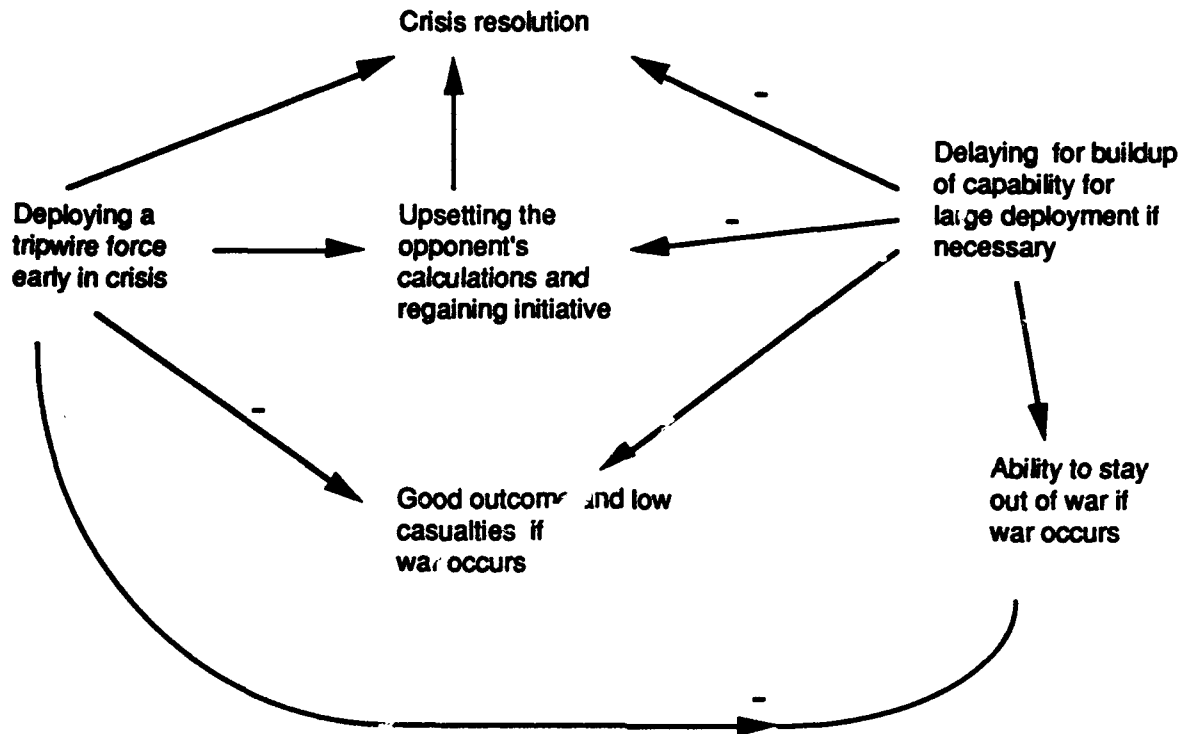


Fig. 2.5—An Illustrative Influence Diagram

DECISION TABLES

A key element in our approach is the extensive use of judgment, outcome, and decision tables (all of which we will sometimes refer to as decision tables in what follows). These tables have a common structure that allows one to see the whole class of input variables to an overall judgment or decision and to follow the logic of that judgment or decision. Table 2.1 shows a simple judgment table. The table makes use of qualitative variables with values as follows for a simple case:

Prospects: (Bad, Marginal, Good)

Risks (downside potential): (High, Moderate, Low)

Opportunities (upside potential): (Poor, Moderate, Good)

Net assessment: (Bad, Marginal, Good)

For example, the variable **Prospects** has values **Bad**, **Marginal**, and **Good**. In each of these cases, the values are ordered from worst to best.

One can now read the first line of the table's body as

If Prospects are Good and
 Risks are Low and
 Opportunities are Good
Then
 Net Assessment is Good.

The second line corresponds to

Else [i.e., if the first line's conditions were not met]
If Prospects are Good and
 Risks are Low and
 Opportunities are Moderate
Then [note the "/"]
 Net Assessment is Good.

The table format is in some respects akin to a decision tree, but in practice it is more powerful.¹⁸ We have drawn heavy lines above and below one line of the table to show how, in a given situation, one can indicate which of a table's rules applies. In this case, the net assessment is Marginal. Trees become hopelessly complex as the number of cases increases, especially with large numbers of variables. In such cases, the tables also become large. However, if we exploit the fact that the variables have been defined so that "bad" values are to the left of and lower than "good" values, then we can simplify the depiction as illustrated in Table 2.2, which is logically identical to Table 2.1. Twenty-seven lines become seven. In our experience, communication of concepts is often easier with the more compact table, despite its increased sophistication.

Tables 2.1 and 2.2 might represent a net assessment of continuing current strategy or implementing a new one. Thus, for each alternative strategy there could be an assessment table of this form. Presumably, the decisionmaker will choose the option with the most favorable net assessment. Table 2.3 is an options-comparison table that one can construct for a given situation (and for a given decisionmaker). In other situations, the options might compare quite differently. Thus, Tables 2.1 and 2.2 are useful for analyzing the goodness of options under a range of situations, while Table 2.3 is useful in comparing across options for a single situation. Table 2.3 is especially useful in group discussions. Note that if some

¹⁸Formal computer models can be developed using tables almost precisely like those shown here, if one uses the RAND-ABEL[®] programming language. Such models are described, for example, in Davis, Banks, and Kahan (1986). The language is described in Shapiro, Hall, Anderson, LaCase, Gillogly, and Weissler (1988) and Davis (1990).

Table 2.1
An Illustrative Judgment Table
(for a risk-averse individual)

Prospects (best-estimate future) ^a	Risks (downside potential or worst-case future)	Opportunities (upside potential or best-case future)	/Net Assessment of /Situation and Trends ^b
=====	=====	=====	=====
Good	Low	Good	Good
Good	Low	Moderate ^{*c}	Good*
Good	Low	Poor*	Marginal*
Good	Moderate	Good	Good
Good	Moderate	Moderate*	Marginal*
Good	Moderate	Poor*	Marginal*
Good	High	Good	Bad
Good	High	Moderate*	Bad*
Good	High	Poor*	Bad*
Marginal	Low	Good	Good
Marginal	Low	Moderate	Marginal
Marginal	Low	Poor*	Marginal
Marginal	Moderate	Good	Bad
Marginal	Moderate	Moderate	Bad
Marginal	Moderate	Poor*	Bad*
Marginal	High	Good	Bad
Marginal	High	Moderate	Bad
Marginal	High	Poor*	Bad*
Bad	Low*	Good	Bad*
Bad	Low*	Moderate	Bad*
Bad	Low*	Poor*	Bad*
Bad	Moderate*	Good	Bad*
Bad	Moderate*	Moderate	Bad*
Bad	Moderate*	Poor	Bad*
Bad	High	Good	Bad
Bad	High	Moderate	Bad
Bad	High	Poor	Bad

^aValues of variables are as follows: Prospects: (Bad, Marginal, Good); Risks: (High, Moderate, Low); Opportunity: (Poor, Moderate, Good); Net Assessment: (Bad, Marginal, Good).

^bThe logic used here assumes that High Risks imply a net assessment of Bad; if Prospects are only Marginal, then net assessment is Bad if Risks are Moderate or High. Other calculations "average" the three assessments. Other definitions of risk aversion are obviously possible, but this definition emphasizes thresholding.

^cThe * denotes a logically impossible case such as one having Opportunity (best-case outcome) being worse than Prospects (the best-estimate outcome).

options have the same first-order net assessment, the decisionmaker may tilt toward one or another on grounds such as the desire to "do something," or, with contrary reasoning, to "not upset the applecart." Whatever those considerations are (e.g., risk-accepting behavior motivated by an "intolerable" current situation, a sense of holding the initiative, and conqueror attitudes), they can be represented by additional columns so that there will always be a single "best" option except if one wants to reflect random tie breakers.

Table 2.2
Efficient Representation of Same Judgment Table

Prospects =====	Risks ^a =====	Opportunities =====	/Net Assessment of /Situation and Trends =====
Good	≥Moderate	≥Moderate	Good
Good	≥Moderate	Poor*	Marginal*
Good	High	≥Poor	Bad
Marginal	Low	≥Moderate	Marginal*
Marginal	Low	Poor*	Marginal*
Marginal	≤Moderate	≥Poor	Bad
Bad	≥High*	≥Poor	Bad*

^aNote that the values of Risks are ordered as High, Moderate, Low, with High < Low in the sense that high risks are worse than low risks.

Table 2.3
An Options-Comparison Table
(for a specific situation and decisionmaker)

Option =====	Prospects under Option =====	Risks under Option =====	Opportunities under Option =====	/Net Assessment /of Option =====
1	Good	Low	Good	Very Good
2	Good	High	Poor	Marginal
3	Marginal	Low	Poor	Marginal

Those familiar with standard methods of decision analysis may be puzzled (or even put off) by this decision-table format with its use of qualitative variables. However, we believe it is cognitively natural as a description of the way people effectively balance different considerations. Further, we have found it to be highly successful for discussing decisionmaking at the analyst-to-analyst level and, importantly, in groups and lectures. Appendix B compares this approach to more standard decision analysis methods.

III. APPLYING THE METHODOLOGY

In this section we sketch how one applies the generic methodology to a particular case. We assume that the application is contingency planning for a possible crisis involving the country Alpha and the dictator Zed.

BACKGROUND STAFF WORK

Preliminary Descriptions

Clearly, no serious work could be accomplished without regional experts collecting various elements of basic information such as where Alpha is geographically, who its neighbors are, on what its economy depends, the nature of its political system, its historical relationships with neighbors, and so on. Of particular interest here would be historical disputes that might lead to crisis—e.g., disputes over territory, rights of passage, fishing areas, cross-border activities by political activists, terrorism, and religion. Relatively few crises emerge “from nowhere” and revolve around issues that have not been visible in advance. A possible Iraqi invasion of Kuwait, for example, was considered time and again by strategists in the decade before August, 1990.

Describing Perspectives and Requiring Multiple Models

After collecting appropriate objective descriptions, attention should begin shifting to identifying and understanding various ideas and perspectives. This, of course, can be difficult when dealing with a region far away with a history and culture very different from our own. Regional specialists are again essential—preferably ones knowing the relevant languages and with experience living in the society, not just within embassy walls.

We have no formulas for how to accomplish this, because the difficulties are notorious. For example, it is all too easy for us to believe that what a nation's educated elite is telling us is representative of the ideas in the street. Arguably, that is what happened with respect to Iran, where, reportedly, the United States had foregone emphasizing human intelligence collection out of deference to the Shah. Despite the difficulties, however, there is every reason to believe that regional specialists can succeed if given a chance. There is one important guideline here, however: *regional specialists should be asked to describe all the players, ideas, factors, and possibilities (including, importantly, ones they do not regard as likely); they should not be relied upon for high-confidence predictions and should not be encouraged to make them lightly.* Further, in characterizing the thinking of Zed himself (or,

better, his behavior in a context that reflects his advisors, political constraints, etc.), *the specialists should sketch at least two plausible "models."* Given the extreme difficulty we have predicting political events in our own country, why should we ask regional specialists for high-confidence predictions about foreign countries or their leaders?

There is more to this guideline than an innocuous admonition. The essence of our methodology is that one can have much higher confidence in our ability to identify the factors that would influence a nation's decisions than our ability to predict what those decisions will be, except in limiting cases. Regional specialists are often superb at identifying the factors and at describing them with care and subtlety. By contrast, they have an understandably abominable record in predicting crises.¹⁹ Figure 3.1 illustrates this with a list of past shocks—i.e., crises or events that had not been predicted, at least not effectively.

What are some of the ideas and perspectives that regional specialists might bring to the attention of contingency planners? Drawing on Fig. 3.1, we think of the following—merely as examples. In each case the date shows when the statement that follows it might reasonably have been made as part of contingency planning. By "might reasonably have



Fig. 3.1—An Exercise in Humility: Past Shocks

¹⁹There are parallels to this in other fields. For example, Dawes, Faust, and Meehl (1989) provide empirical evidence on how clinical specialists are often much better at identifying key variables than in developing coherent explanatory models (diagnoses).

been made" we mean that under circumstances that encouraged regional specialists to lay out a range of ideas and possibilities, rather than mere best-estimate judgments, we believe the regional specialists might well have made these statements.²⁰ In some cases, they probably did.

- [1978] Despite the apparent stability of the Shah's regime, there is strong grass-roots resentment, especially among the Shia population. The resentment is about everything Western, with many people finding the idea of a fundamentalist Islamic state attractive. Although fundamentalist Muslims vary significantly, the most esteemed current figure in Iranian circles is Ayotollah Khomeini, residing in Paris currently. He has a deep hatred for the West (and believes the death of one of his sons was the result of CIA actions). He tends to believe in Western conspiracies against Iran and emphasizes the fact that the CIA installed the Shah.
- [1978] In a crisis with Iranian Islamic fundamentalists, the United States might see the opponent taking hostages, since there is a long tradition of that in Middle Eastern history. Taking hostages is not considered "wrong," as it would be in the West.
- [1972-1973] President Sadat of Egypt does not have the passionate anti-Israeli hatred that characterizes some of the other Middle Eastern leaders, but he is in a sense spoiling for a fight. There is a view that Egypt cannot turn inward until and unless it regains some honor by successfully dealing with Israel in some military crisis. If war did erupt, Sadat would not necessarily insist on a clear-cut military victory, unless it were clearly achievable, which seems unlikely.
- [1982] The attitudes of Iraq and Iran toward the use of chemical weapons are probably quite different from those in the West. They lack nuclear weapons but are engaged in an all-out war. Should either of them begin to lose catastrophically, we must assume that it would consider using chemicals—with at least as much logic as we assume that NATO would use nuclear weapons if losing a war with the Soviet Union. Should Iraq and Iran employ chemical weapons or ballistic missiles, it might be specifically as instruments of terror rather than warfighting.

This list includes nothing from the recent war with Iraq over Kuwait. We discuss that conflict in some detail in a companion paper (Davis and Arquilla, 1991). It is worth noting here, however, that a contingency planning workshop held by RAND in mid-July 1990 had as its first case a possible Iraqi threat to Kuwait and Saudi Arabia. By adopting the approach we suggest—urging specialists to play through the *possibilities* rather than focusing on the best estimate—it was possible to identify virtually all of the "moves." The workshop

²⁰The word *might* is important here because of the well-known and insidious effects of hindsight, which include the tendency to believe that we *would* have predicted accurately, if only See especially Hawkins and Hastie (1990).

participants did not expect an invasion, but they worked through the problem anyway and developed an excellent sense for the gameboard.²¹

Identifying Key Variables and Developing Hierarchies

The next step is to identify and name key variables (which some may think of as factors rather than variables), proceeding in a top-down way so that one first identifies the handful of key top-level variables (e.g., Risks) and only then breaks them down into second-, third-, and even fourth-level variables. Naming the variables well is critical if there is to be effective group discussion later.

Developing Influence Diagrams and Cognitive Maps

If variables are the bricks of the house one is building, then influence diagrams or cognitive maps are the rough sketches that provide alternative perspectives of what one wants the building to look like. If our eventual model is to be natural for the purpose intended, it must not only be expressed in the natural variables of the opponent's perspective, it must also reflect the cause-effect relationships that he perceives as most important. So, for example, an appropriate cognitive map might show assumed conspiracies between the United States and its allies acting against the economic or political interests of Alpha—whether or not such a conspiracy exists. Since specialists will disagree about these cognitive maps, there should be alternatives developed.

After developing cognitive maps, it should be possible to develop simplified diagrams (such as the one in Fig. 2.4) showing the most important variables affecting particular judgments (e.g., the assessment of alternative courses of action) and the relative weight of those variables.

Developing First-Cut Decision Tables for Different Decision Models of the Decisionmaker

With the background of diagram development, it is relatively straightforward to build decision tables of various kinds, again starting from the top and working downward only as far as seems necessary. Often, in our experience, the top two levels are quite adequate to guide considerable analysis and discussion.²²

²¹One of us (Davis) and colleague Zalmay Khalilzad were able, as the result of the workshop, to write and distribute on the day of invasion a strategic overview providing a rather accurate picture of stakes and options for both sides. This and other papers written before or during the crisis will be published as part of a larger report by Khalilzad and Davis.

²²The combining rules themselves are not as straightforward as they at first appear. As illustrated in Table 2.1, we believe descriptive rules apply thresholding in what decision analysts might regard as dubious ways. For example, risky options may not be considered at all when decisionmakers

GROUP DISCUSSION

Armed with the posited staff work, the next step might be group discussions—first with other staff and later with senior figures. Ideally, the group would include people with a mix of backgrounds and perspectives. A traditional presentation with slides and an overhead projector is workable, as is a blackboard approach. Increasingly, however, RAND is using a Macintosh computer with a large-screen projector.

The briefer should be highly substantive and senior enough to engage directly in debate—without having to submit to authority readily when someone in the group disputes an argument. He should also be open and flexible enough to extract from the group discussion the key elements that should be added—in real time—to the charts. This might be done with the low-technology approach of a grease pencil or with the increasingly common procedure of editing the slide with a microcomputer (e.g., editing Power Point slides with a Macintosh computer). There are some limitations (e.g., few groups will have multiple large screens and large-screen projectors), so having auxiliary blackboards or large tablets handy can be very useful.

In our experience, having the strawman structure and model is crucial to discussion. If the strawman is good, then the audience will almost instantly begin identifying key issues on which there are valid disputes. Further, the disputes will be rather well defined, with relatively little “talking past one another,” primarily because everyone can see “all” the variables at the same time and everyone is using the same frame of reference. Obviously, poor staff work leading to a poor strawman will lead to a poor group discussion: the group will “tear up” the basic frame of reference.

ANALYSIS

Analysis, in the context of our approach and model, includes (a) defining variables and their values in some detail, addressing subtleties of meaning; (b) developing careful and suitable hierarchies of variables, recognizing that many different hierarchies are possible but that some will prove more natural than others; (c) refining names of variables and their values so that they become more nearly self-evident in meaning; (d) working out the decision tables—covering all cases, rather than merely those of greatest perceived importance; (e) restructuring the tables to be efficiently comprehensible.

Analysis should be conducted both before and after group discussions. It is needed beforehand to assure that the strawman model is sound enough to guide the group

are risk averse. By contrast, “smart” decisionmaking might balance risks against gains in an explicit calculation. Combining rules are discussed further in Davis and Arquilla (1991), but further research is appropriate.

discussion. It is needed afterward to accommodate the new insights and differences of opinion and to fill in the enormous gaps: it is, after all, impractical to expect a group to discuss more than a few of the many possible combinations of variables. Whether the group discussion provides the necessary guidance for subsequent analysis purporting to represent the group in some respect is highly dependent on the individuals involved and many other factors. Iteration is essential.

AMENDING ONE'S OWN STRATEGY

No matter how diligently one works on opponent modeling, it is unlikely that the results will be compellingly predictive. There are simply too many unknowns and random factors. The opponent models can be extremely useful, nonetheless, if they clarify "the strings to pull and the levers to push." The objective, after all, is to develop a well-hedged strategy that will influence the opponent toward thinking, or at least behaving, in a way favorable to our own interests. What actions should we and our allies take (or avoid) with this in mind? Would they be sufficiently likely to be successful so that we could accept the downside risks associated with them? These are the kinds of questions one wants to ask with the benefit of opponent models as background.

In subsequent work we shall discuss these issues in more detail, but here it may be sufficient merely to itemize some *examples* of what one might do as the result of such an analysis:

- Send an emissary or a third party to lay out in detail implications that the opponent may have overlooked or misunderstood: spell out his risks under aggression and his benefits under more benign circumstances.
- Undercut his assumptions about our own thinking and likely actions, and those of third parties, and about the likely results of combat (i.e., about the nature and magnitude of damage and his ability to recover).
- Undercut his faith in his ability to maintain the initiative and control events. Sensitize him to uncertainty. Demonstrate one's own resolve (and, perhaps, one's unpredictability).²³
- Lay out and discuss options that he may not have adequately considered. In some cases, calm his fears about bad consequences of options favorable to one's own interests (perhaps providing credible assurances)

²³The opponent, of course, has his own mental model of us. One function of strategy should be to change that mental image. Taking this farther, his mental model of us may include what he thinks our mental model of him is. That also is something we might want to change. This process could in theory be applied recursively on to infinity. In practice, that is unnecessary. For related discussion in the context of U.S. and Soviet models of each other, see Davis, Bankes, and Kahan (1986) and subsequent analysis. See also the reflexive-control work of Lefebvre and Lefebvre (1984 and 1985).

Most of these should be thoroughly familiar admonitions for those who have studied the literature of crisis or negotiations, or who merely have a good intuition for bargaining. However, it is only when one inserts details that they become important, and getting the details right depends on getting "the picture" (or at least one of the possible pictures taken seriously) more nearly right than wrong. Hence, the concept of opponent modeling.

Appendix A

CONDITIONAL RATIONALITY

Rationality is exceptionally difficult to define, and there is little agreement to be found in the literature. To illustrate this and to better describe what we have in mind by *limited rationality*, consider the following as our effort to describe and compare common definitions of rationality held explicitly or implicitly by various individuals.

1. *Decision Analysis.* Rationality requires considering the utilities of various outcomes, the probability of those outcomes for each option, and a calculation such as how to maximize expected utility. The utilities must satisfy the transitivity principle (If A is better than B, and B is better than C, then A must be better than C).
2. *Decision Analysis Plus Search.* Rationality requires decision-analytic thinking preceded by a reasonable search for adequate options.²⁴
3. *Systems Analysis.* Good decisionmaking requires formulating a diverse set of strategies, comparing those strategies in various analytic ways (e.g., estimating effectiveness, in several dimensions, of equal-cost alternatives), and considering the results in a framework allowing for the qualitative application of values, judgments, and intuition to factors not readily treated analytically.
4. *Cognitive Psychology.* Rational decisionmaking requires not only the methods of decision analysis where possible, but also special measures to avoid standard psychological blunders such as seeing only organizationally blessed options, groupthink, sequential rather than parallel assessment of options, underestimating risks, etc.
5. *Limited Rationality (as used in this paper).* Rational decisionmaking requires a fairly good set of options ("reasonable" search), parallel examination of options with an eye on objectives, and explicit consideration of upside and downside potential as well as likely outcomes. It does *not* require formulating anything like utility functions in doing so. Decisions made under limited rationality will have a superficially logical basis—i.e., a "reasonable" relationship between objectives and decisions.²⁵ They may, however, suffer from a wide range of errors and misperceptions. These include inappropriate framing; thresholding, with the

²⁴See Schelling (1980) and Bueno de Mesquita (1981) for related discussions.

²⁵Our models demonstrating limited rationality will satisfy the transitivity principle in any single decision. However, because we explicitly consider how they view the current situation and trends—i.e., because we allow the anchor point to shift over time as a function of events—the models can behave in ways that appear to violate transitivity. To put it differently, the models can "change their minds" about the relative goodness of options as circumstances change their view of the world and their baseline prospects. Further, we emphasize that a real decisionmaker might shift from behavior described well by one model to behavior described well by another as a result, for example, of being shocked by events. Indeed, one of us has argued elsewhere that the concept of flexible response and its associated concept of deliberate nuclear escalation were fatally flawed except if applied in certain circumstances where a NATO nuclear response would so shock a Soviet invader of Europe as to cause him to change anchor points substantially (Davis, 1989c).

effect of eliminating what are perceived to be low-probability outcomes; misperceptions; groupthink; etc. Importantly, value judgments are often made implicitly in limited rational decisionmaking. For example, a risk-acceptant decisionmaker may look at likely, worst-case, and best-case outcomes of an option and focus in on the likely and best-case outcomes—unless the worst-case outcome seems too likely to be discounted.

By declaring someone's behavior as having "limited rationality," then, we do not mean to imply that the decisions are wise or sound. Rather, we mean only that the decisionmaking has satisfied certain minimum criteria for what passes in the real world as rationality.

Appendix B

COMPARISON WITH STANDARD DECISION ANALYSIS

The mapping between our decision tables and the calculations of standard decision analysis is not as straightforward as one might at first think—which is, in itself, an indication of how subtle human reasoning and its limitations are. Having qualitative variables and quantitative variables is *not* the key difference, because one might establish a simple mapping such as:

Very Bad = 0; Bad = 2; Marginal = 4; Good = 6; Very Good = 8

The numbers would represent “utility values.”

One might also map the several individual assessments into something like probabilities. For example, if the subjective probability distribution for utility of outcomes were well-behaved (e.g., Gaussian), one might order the outcomes from worst to best according to their utility along the x axis. The y axis would then be the subjective probability (or probability density) of a given outcome. Then, one could make a mapping such as the one shown in Fig. B.1 and say:

The utility ascribed to what is called “Prospects” or the “Best-Estimate Future” is the average utility of outcomes over the range of outcomes centered at the most probable outcome, which have a combined probability of 50%. The utilities ascribed to “Risks” and “Opportunities” are the average utilities over the ranges to the left and right of the central range.

With such a mapping, one could translate back and forth between our approach and decision analysis, in which one chooses the option that maximizes expected utility U_i where u_{ij} is the utility of outcome j of strategy i and P_{ij} is the probability of that outcome under strategy i :

$$U_i = \sum_j u_{ij}P_{ij}$$

In fact, we do *not* assume any such mapping to decision analysis—nor any such well-behaved probability distribution as that shown in Fig. B.1. Further, in some cases the logic we ascribe to particular types of opponents in particular contexts is highly “nonrational” in

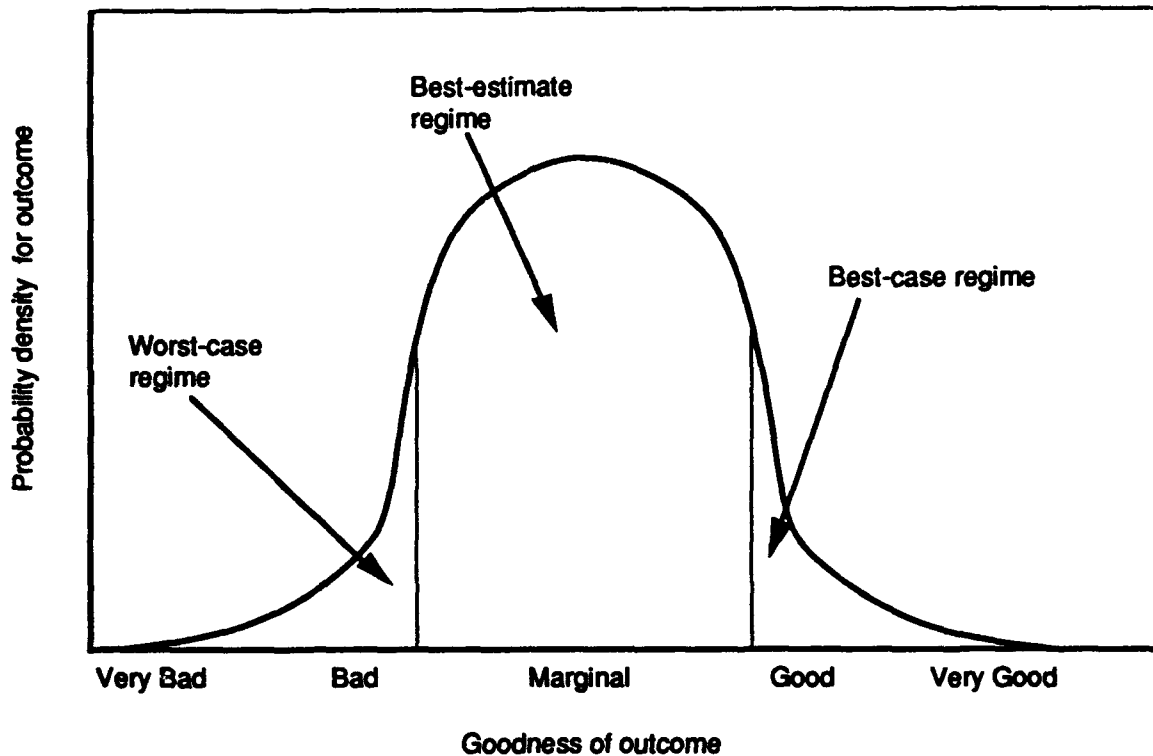


Fig. B.1—Special Case in Which Outcomes and Probabilities Map Easily into Constructs of Current Model

the sense just discussed. For example, a decisionmaker who is personally very dissatisfied with the way things are going may act as though he considered a continuation of the current situation as having infinite disutility—ranking the do-nothing option as Very Bad even though other options involve risks of outcomes that “should” be considered even more unacceptable.

There are many other differences in detail. For example, the rules of decision analysis mathematics result in consistency. By contrast, some of the decision logic we incorporate reflects the inconsistencies real people exhibit routinely. The basic principle here is that our model assumes that people fold together different considerations by a combination of thresholding, averaging, and likelihood maximizing. In some cases they ignore risks (which is typically a combination of underestimating the likelihood of the bad outcomes and of underestimating how bad the bad outcomes are). In other cases they avoid them to an extreme. In looking at options, they may think in terms of picking the option that maximizes the likelihood of the best outcome rather than picking the option that maximizes expected outcome. This is illustrated by Table B.1, which posits that under Options A and B it is most

likely (50 percent) that the outcome will be "marginal" (e.g., much like the current situation). Option A has high risks but the potential for a big improvement. Option B has low risks and no potential. The so-called risk taker, if dissatisfied with the status quo and strongly goal oriented, would choose Option A.

Table B.1
A Possible Risk Taker's Assessment of Options

Option	Prospects	Risks	Opportunity	/Net Assessment
A	Marginal	High	Good	Good
B	Marginal	Low	Poor	Bad

There are many complications in this type of discussion. A decision theorist might look at this assessment of the options and conclude that it is mischaracterized: "In reality, the decisionmaker clearly is giving very high disutility to the current situation; it is therefore inappropriate to describe it as 'Marginal'; it should instead be something like 'Very Bad.'" The same theorist might argue that the decisionmaker is underestimating just how bad the bad outcome associated with Risks is, or that he apparently misperceives the probability of the bad outcome to be lower than it really is. It is certainly the case that one can study the choices made by individuals and infer from them various possible utility functions and, stretching things even further, implicit imputed probabilities. As the behavioral literature demonstrates, however, the bottom line is that people do not reason according to the rules of economic rationality.

We believe, moreover, that this after-the-fact inference about "real" utility functions is not especially helpful in trying to "get into the head" of possible opponents. For example, if someone had asked Hitler about prospects for Germany before he invaded Poland (i.e., in the context of assuming that things would continue as they were rather than with an invasion), it is not at all clear that he would have reported prospects as Very Bad. He might, or he might not. In some contexts, he might have reported on the improved economy and good political trends. In other contexts, he might have argued passionately that it was intolerable for the current borders to be continued and that it was essential that Germany achieve its (his) aspirations. The point here is that Hitler's visceral drive to conquer is not well captured in the language and attitudes of decision analysis—i.e., *decision analysis is not the "natural representation" for descriptive models of decisionmaking and behavior.*

What, then, are the definitions of the variables and values that we use in our decision tables? What instructions *should* we use in asking people to make their assessments or

guess the assessments of potential opponents in crisis or conflict? The answer is that in providing a value for a variable like Prospects we are providing a combined sense of utility and probability—something like a product of the two, but not so neat. Further, in folding together various considerations, the logic we assume can follow any of several styles, which need not be rational analytic, but may be, for example, risk taking, opportunity maximizing, risk avoiding, and so on.

It is crucial to emphasize that this modeling approach is motivated by the desire to describe rather than the desire to prescribe. Real human beings do not process information according to the rational-analytic model. Our model, we believe, does better in this regard. However, one should not be surprised if it contains some inconsistencies—that is part of what we are describing.

That said, we believe our model can be used prescriptively so long as terms are well defined and there is great care in establishing the implicit scales that are to go along with qualitative concepts such as Very Bad, Bad, and so on. This is *not* a trivial matter, because one must constantly distinguish between local and global judgments. That is, when asked whether something is Bad or Marginal, does one think about direct or indirect effects? Further, is the relationship of Very Bad to Bad the same relationship as between Good and Very Good, or between Marginal and Good? That is, are the values “equally spaced” in *some* sense? The qualitative model we are describing here is by no means as simple and intuitive as it may at first appear to be. We have discussed some of these issues in more detail technically elsewhere (e.g., Davis, Bankes, and Weissler, 1989; Davis and Arquilla, 1991, Appendix A).

BIBLIOGRAPHY

- Allison, Graham (1971), *Essence of Decision: Explaining the Cuban Missile Crisis*, Little, Brown, and Co., Boston, MA.
- Arquilla, John (1991), *Theory of Interstate War*, Ph.D. dissertation, Political Science Department, Stanford University, CA.
- Arrow, Kenneth (1963), *Social Choice and Individual Values*, 2d edition, John Wiley, NY. Paperback available from Yale University Press, New Haven, CT.
- Axelrod, Robert (1976), *The Structure of Decision*, Princeton University Press, NJ.
- Brams, Steven J. (1985), *Superpower Games: Applying Game Theory to Superpower Conflict*, Yale University Press, New Haven, CT.
- Bueno De Mesquita, Bruce (1981), *The War Trap*, Yale University Press, New Haven, CT.
- Davis, Paul K. (1987), "A New Analytic Technique for the Study of Deterrence, Escalation Control, and War Termination," in Stephen J. Cimbala (ed.), *Artificial Intelligence and National Security*, D. C. Heath and Company, Lexington, MA.
- Davis, Paul K. (1989a), *National Security Planning in an Era of Uncertainty*, RAND, P-7605. This will appear also as a chapter in *Conventional Forces and Arms Control: Technology and Strategy in a Changing World*, a book stemming from a conference held September 25-26, 1989, at Los Alamos National Laboratory.
- Davis, Paul K. (1989b), *Studying First-Strike Stability with Knowledge-Based Models of Human Decisionmaking*, RAND, R-3689-CC.
- Davis, Paul K. (1989c), *Some Lessons Learned from Building Red Agents in the RAND Strategy Assessment System (RSAS)*, RAND, N-3003-OSD.
- Davis, Paul K. (1990), *An Analyst's Primer for the RAND-ABEL Programming Language*, RAND, N-3042-NA.
- Davis, Paul K., and John Arquilla (1991), *Deterring or Coercing Opponents in Crisis: Lessons from the War with Saddam Hussein*, RAND, R-4111-JS, forthcoming.
- Davis, Paul K., and Barry Wolf (1991), "Behavioral Factors in Nuclear De-Escalation," to be published by St. Martin's Press in a book edited by Joseph Nation.
- Davis, Paul K., and Peter J.E. Stan (1984), *Concepts and Models of Escalation*, RAND, R-3235.
- Davis, Paul K., Steven C. Bankes, and James Kahan (1986), *A New Methodology for Modeling National Command-Level Decisionmaking in War Games and Simulations*, RAND, R-3290-NA.

- Davis, Paul K., Steven C. Bankes, and Robert Weissler (1989), "Formalization of Tradeoff Rules and Other Techniques for Comprehending Complex Rule-Based Models," in Maurice S. Elzas, Tuncer I. Oren, and Bernard P. Zeigler (eds.), *Modelling and Simulation Methodology: Knowledge Systems' Paradigms*, North-Holland, Amsterdam. Also available as RAND P-7394.
- Dawes, Robin, and B. Corrigan (1974), "Linear Models in Decision Making," *Psychological Bulletin*, Vol. 81, pp. 95-106.
- Dawes, Robin, David Faust, and Paul Meehl (1989), "Clinical Versus Actuarial Judgment," *Science*, Vol. 243, pp. 1668-1674.
- Dunning, David, Dale W. Griffin, James Milojkovic, and Lee Ross (1990), "The Overconfidence Effect in Social Prediction," *Journal of Personality and Social Psychology*, Vol. 58, No. 4, pp. 568-581.
- Fischhoff, Baruch (1982), "Debiasing," in Daniel Kaneman, Paul Slovic, and Amos Tversky (eds.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, England.
- George, Alexander, and Richard Smoke (1974), *Deterrence in American Foreign Policy: Theory and Practice*, Columbia Press, NY.
- Hawkins, Scott A., and Reid Hastie (1990), "Hindsight: Biased Judgments of Past Events After the Outcomes Are Known," *Psychological Bulletin*, Vol. 107, pp. 311-327.
- Holsti, Ole (1989), "Crisis Decision Making," in Philip E. Tetlock, Jo L. Husbands, Robert Jervis, Paul C. Stern, and Charles Tilly (eds.), *Behavior, Society, and Nuclear War, Vol. I*, Oxford University Press, NY.
- Huth, Paul (1988), *Extended Deterrence and the Prevention of Wars*, Yale University Press, New Haven, CT.
- Huth, Paul, and Bruce Russett (1988), "Deterrence Failure and Crisis Escalation," *International Studies Quarterly*, Vol. 32, pp 29-45.
- Isenberg, Daniel J. (1986), "Group Polarization: A Critical Review and Meta-Analysis," *Journal of Personality and Social Psychology*, Vol. 50, pp. 1141-1151.
- Janis, I. L., and L. Mann (1977), *Decision-Making*, Free Press, NY.
- Janis, Irwin (1982), *Groupthink*, 2d edition, Houghton Mifflin, Boston, MA.
- Jervis, Robert (1976), *Perceptions and Misperceptions in International Politics*, Princeton University Press, NJ.
- Kahneman, Daniel, and Dan Lovallo, "Timid Decisions and Bold Forecasts: A Cognitive Perspective on Risk Taking," paper presented at the Conference on Fundamental Issues in Strategy, Silverado, November 29-December 1, 1990.
- Kahneman, Daniel, and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, Vol. 47, pp. 263-291.

- Kahneman, Daniel, Paul Slovic, and Amos Tversky (eds.) (1982), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, England.
- Kanwisher, Nancy (1989), "Cognitive Heuristics and American Security Policy," *Journal of Conflict Resolution*, Vol. 33, No. 4, pp. 652-675.
- Karni, E. (1990), "Preference Reversals," in J. Eatwell, M. Milgate, and P. Newman (eds.), *The New Palgrave: Utility and Probability*, 1st American edition, W. W. Norton & Co., NY.
- Kull, Steven (1988), *Minds at War: Nuclear Reality and the Inner Conflicts of Defense Policymakers*, Basic Books, NY.
- Lefebvre, Vladimir, and Victorina Lefebvre (1984 and 1985), *Reflexive Control*, Science Applications Inc., SAIC-84-024-FSRC-E and SAIC-85/6016&FSRC/E, Greenwood Village, CO.
- Lichtenstein, S., and Paul Slovic (1971), "Reversal of Preferences Between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology*, Vol. 89, pp. 46-55.
- March, James G., and Roger Weissinger-Baylon (1986), *Ambiguity and Command: Organizational Perspectives on Military Decision Making*, Pitman, Marshfield, MA.
- Myers, David G., and Helmut Lamm (1976), "The Group Polarization Phenomenon," *Psychological Bulletin*, Vol. 83, pp. 602-627.
- Platt, J. R. (1964), "Strong Inference," *Science*, Vol. 146, pp. 347-353.
- Ross, L., and C. A. Anderson (1982), "Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments," in Kahneman, Slovic, and Tversky, 1982.
- Ross, Lee, and Richard E. Nisbett (1990), *The Person and the Situation: Perspectives of Social Psychology*, McGraw-Hill, NY.
- Schelling, Thomas C. (1980), *The Strategy of Conflict*, 2d edition, Harvard University, Cambridge, MA.
- Schwenk, Charles R. (1990), "Effects of Devil's Advocacy and Dialectical Inquiry on Decision Making: A Meta-Analysis," *Organizational Behavior and Human Decision Processes*, Vol. 47, pp. 161-176.
- Shapiro, Norman Z., H. Edward Hall, Robert H. Anderson, Mark L. LaCasse, Marrietta S. Gillogly, and Robert E. Weissler (1988), *The RAND-ABEL Programming Language: Reference Manual*, RAND, N-2367-1-NA.
- Simon, Herbert A. (1982), *Models of Bounded Rationality*, MIT Press, Cambridge, MA.
- Steinbruner, John (1974), *The Cybernetic Theory of Decision: New Dimensions of Political Analysis*, Princeton University Press, NJ.

- Tetlock, P. E. (1988), "Monitoring the Integrative Complexity of American and Soviet Policy Rhetoric: What Can Be Learned?" *Journal of Social Issues*, Vol. 44, pp. 101-132.
- Tversky, Amos, and Daniel Kahneman (1981), "The Framing of Decisions and the Rationality of Choice," *Science*, Vol. 211, pp. 453-458.
- Tversky, Amos, and Daniel Kahneman (1982), "Causal Schemas in Judgments Under Uncertainty," in Kahneman, Slovic, and Tversky, 1982.
- Vallone, Robert P., Dale Griffin, Sabrina Lin, and Lee Ross (1990), "Overconfident Prediction of Future Actions and Outcomes by Self and Others," *Journal of Personality and Social Psychology*, Vol. 58, No. 4, pp. 582-592.
- Whyte, G. (1989), "Groupthink Reconsidered," *Academy of Management Review*, Vol. 14, pp. 40-56.
- Winnefeld, James A., and David A. Shlapak (1990), *The Challenge of Future Nonstandard Contingencies: Implications for Strategy, Planning, Crisis Management, and Forces*, Vols. I and II, RAND, N-3098/1-DAG and N-3098/2-DAG.